


<b>Nama:</b> <b>Ibnu Fajar Setiawan</b>  <b>NIM:</b> <b>065002000006</b>	 <b>Praktikum Data Warehouse</b>	<b>MODUL 6</b>  <b>Nama Dosen:</b> <b>Ir. Teddy Siswanto, MMSi</b>
<b>Hari/Tanggal:</b> <b>Hari, 19/05/2022</b>		<b>Nama Asisten Labratorium:</b> <b>1. Azhar Rizki Zulma</b> <b>065001900001</b> <b>2. Nadiya Amanda Rizkania</b> <b>064001900003</b>

## Merge Data pada Spoon

### 1. Teori Singkat

Data warehouse adalah jenis sistem manajemen data yang dirancang untuk memungkinkan dan mendukung kegiatan business intelligence (BI), terutama analitik. Gudang data semata-mata dimaksudkan untuk melakukan kueri dan analisis dan sering berisi sejumlah besar data historis. Data dalam gudang data biasanya berasal dari berbagai sumber seperti file log aplikasi dan aplikasi transaksi. Gudang data memusatkan dan mengkonsolidasikan sejumlah besar data dari berbagai sumber. Kemampuan analitisnya memungkinkan organisasi untuk memperoleh wawasan bisnis yang berharga dari data mereka untuk meningkatkan pengambilan keputusan. Seiring waktu, ia membangun catatan sejarah yang dapat sangat berharga bagi para ilmuwan data dan analis bisnis. Karena kemampuan ini, gudang data dapat dianggap sebagai "sumber kebenaran tunggal" organisasi.

### 2. Alat dan Bahan

Hardware : Laptop/PC

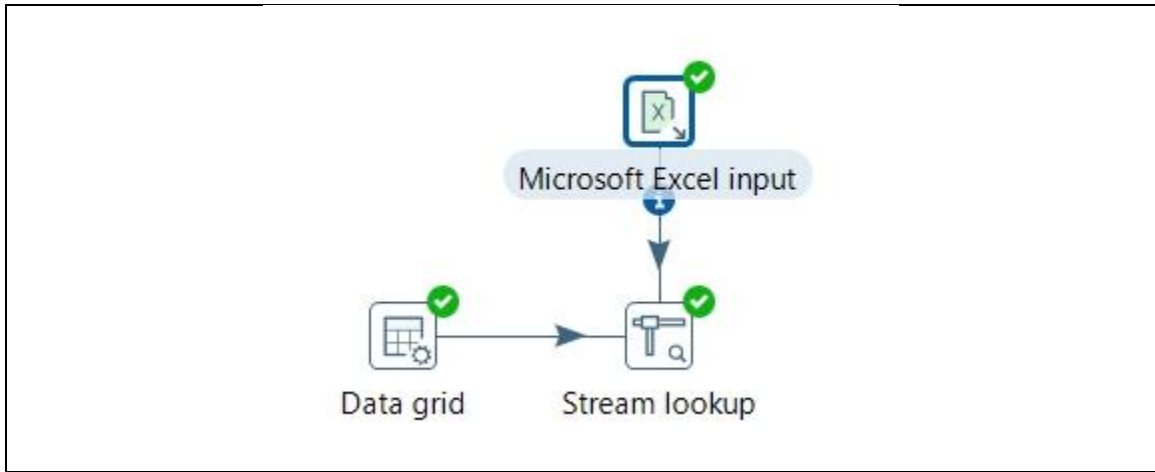
Software : Spoon Pentaho from Hitachi Vantara



### 3. Elemen Kompetensi

a. Latihan pertama – Melakukan penggabungan data menggunakan Stream Lookup

1. Buat transformation baru dengan struktur seperti pada gambar dibawah ini.



2. Microsoft Excel Input.



**Jurusan Teknik Informatika & Sistem Informasi**  
Fakultas Teknologi Industri – Universitas Trisakti

Microsoft Excel input

Step name: Microsoft Excel input

Files | Sheets | Content | Error Handling | Fields | Additional output fields

#	Name	Type	Length	Precision	Trim type	Repeat	Format	Currency	Decimal	Grouping
1	NIS	String	-1	-1	none	N				
2	Nama	String	-1	-1	none	N				
3	Jenis Kelamin	String	-1	-1	none	N				
4	Kelas	String	-1	-1	none	N				

Get fields from header row...

Help OK Preview rows Cancel

### 3. Data Grid.

Data grid

Step name: Data grid

Meta | Data

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Null if	Set empty string?
1	Nama	String								N
2	Kelas	String								N

Help OK Preview Cancel

Data grid

Step name: Data grid

Meta | Data

#	Nama	Kelas
1	MAWAR MELATI	A
2	MAWAR MELATI	B

Help OK Preview Cancel



4. Stream lookup (Get fields, Get lookup fields & hapus variabel yang tidak digunakan).

**Stream lookup**

Step name: Stream lookup

Lookup step: Microsoft Excel input

The key(s) to look up the value(s):

#	Field	LookupField
1	Nama	Nama
2	Kelas	Kelas

Specify the fields to retrieve :

#	Field	New name	Default	Type
1	NIS			String

Preserve memory (costs CPU) ☒

Key and value are exactly one integer field ☐

Use sorted list (i.s.o. hashtable) ☐

Buttons: ? Help, OK, Cancel, Get Fields, Get lookup fields

5. Ketika dijalankan maka outputnya akan seperti gambar berikut ini.

**Execution Results**

Logging | Execution History | Step Metrics | Performance Graph | Metrics | Preview data

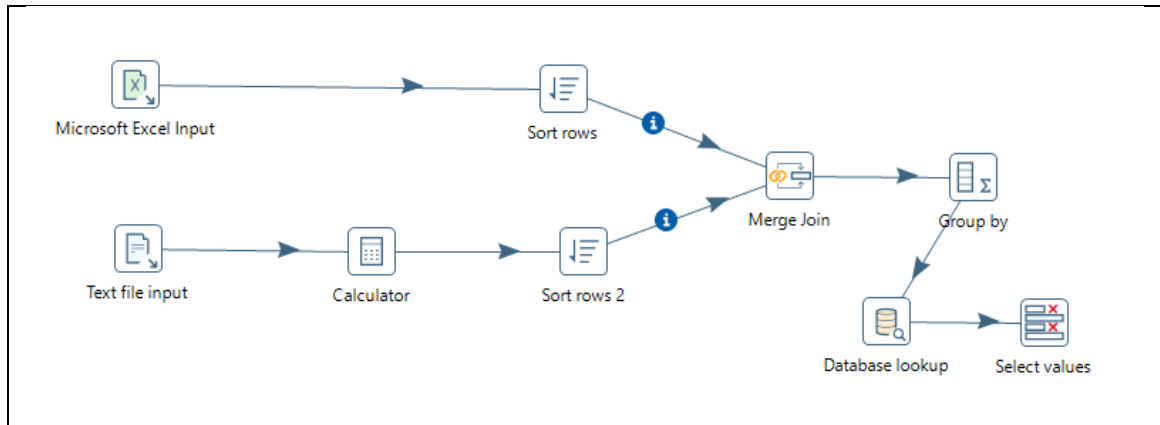
☒ First rows ☐ Last rows ☐ Off

#	Nama	Kelas	NIS
1	MAWAR MELATI	A	<null>
2	MAWAR MELATI	B	<null>

b. Latihan Kedua – Penggabungan dengan Merge dan Database Keylookup

1. Nyalakan Apache dan MySQL pada XAMPP dan buat transformasi baru dengan struktur sebagai berikut.





2. Microsoft Excel Input (Get sheetsname dan Get fields).



② Help

Microsoft Excel input

Step name: Microsoft Excel input

Files | Sheets | Content | Error Handling | Fields | Additional output fields

#	Name	Type	Length	Precision	Trim type	Repeat	Format	Currency	Decimal	Grouping
1	orderNumber	Number	-1	-1	none	N				
2	orderDate	Date	-1	-1	none	N				
3	requiredDate	Date	-1	-1	none	N				
4	shippedDate	Date	-1	-1	none	N				
5	status	String	-1	-1	none	N				
6	customerNumber	Number	-1	-1	none	N				
7	comments	String	-1	-1	none	N				

Get fields from header row...

Help OK Preview rows Cancel

### 3. Text file input (Get fields pada tab fields).



Text file input

Step name: Text file input

File | Content | Error Handling | Filters | Fields | Additional output fields

File or directory:  Add Browse...

Regular Expression:

Exclude Regular Expression:

Selected files:

#	File/Directory	Wildcard (RegExp)	Exclude wildcard	Required	Includ
1	D:\file kuliah smstr4\filepraktadatawarehouse\Dataset (1)\orderdetails.txt			N	N

Delete Edit

Accept filenames from previous steps

Accept filenames from previous step ☐

Pass through fields from previous step ☐

Step to read filenames from:

Field in the input to use as filename:

Show filename(s)... Show file content Show content from first data line

Help OK Preview rows Cancel

---

Text file input

Step name: Text file input

File | Content | Error Handling | Filters | Fields | Additional output fields

#	Name	Type	Format	Position	Length	Precision	Currency	Decimal	Group	Null if	Default	Trim type	Repeat
1	orderNumber	Integer	#		15	0	IDR	.	,	-		none	N
2	productCode	String			9		IDR	.	,	-		none	N
3	quantityOrdered	Integer	#		15	0	IDR	.	,	-		none	N
4	priceEach	Number	#,.		6	2	IDR	.	,	-		none	N
5	orderLineNumber	Integer	#		15	0	IDR	.	,	-		none	N

Get Fields Minimal width

Help OK Preview rows Cancel

#### 4. Calculator.



Calculator

Step name  
Calculator

☒ Throw an error on non existing files

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask	Decimal symbol	Grouping symbol	Currency symbol
1	Total	A * B	quantit...	priceEa...		Number			N	##			

Help

OK Cancel

## 5. Sort rows (Get fields)



[illegible]

6. Sort rows 2 (Get fields 2).

Sort rows

Step name: Sort rows 2

Sort directory: %%java.io.tmpdir%% Browse...

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files? ☒

Only pass unique rows? (verifies keys only) ☐

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted:
1	orderNumber	Y	N	N	2	N

Help OK Cancel Get Fields

## 7. Merge join.

Merge join

Step name: Merge join

First Step: Sort rows

Second Step: Sort rows 2

Join Type: INNER

Keys for 1st step:

#	Key field
1	orderNumber

Get key fields

Keys for 2nd step:

#	Key field
1	orderNumber

Get key fields

Help OK Cancel

## 8. Group by (Get fields & Get lookup fields).



Group by

Step name
Group by

Include all rows?
☐

Temporary files directory
%%java.io.tmpdir%%
Browse...

TMP-file prefix
grp

Add line number, restart in each group
☐

Line number field name

Always give back a result row
☐

The fields that make up the group:

#	Group field
1	orderNumber
2	

Get Fields

Aggregates :

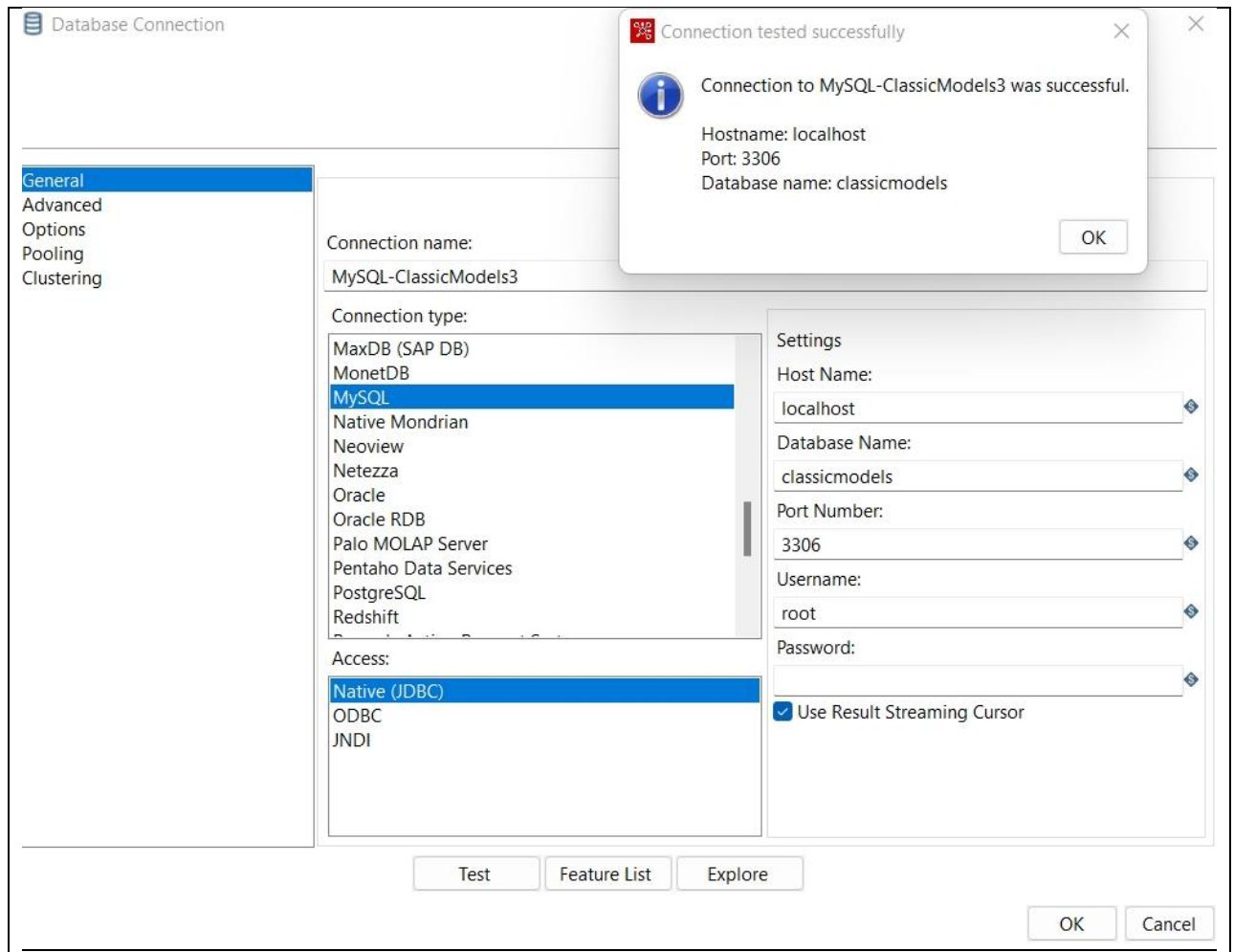
#	Name	Subject	Type
1	customerNumber	customerNumber	First Value
2	Total	Total	Sum

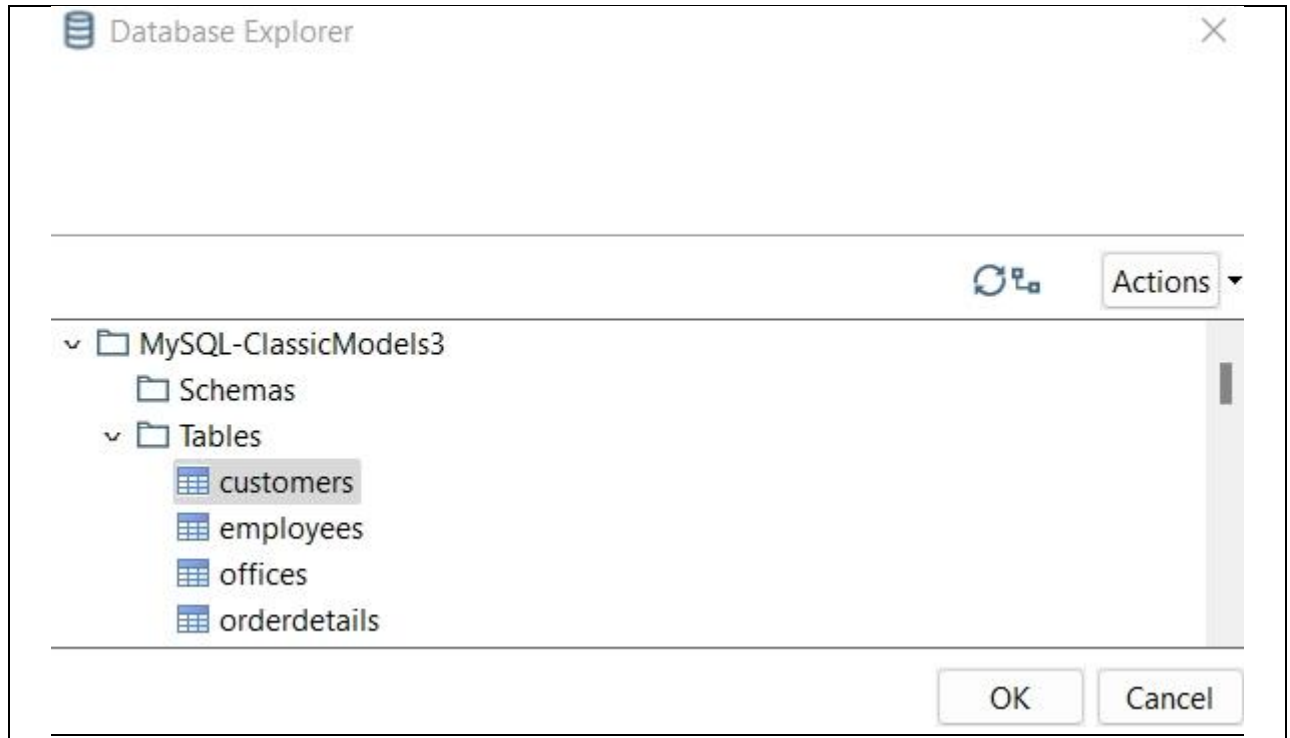
Get lookup fields

Help
OK
Cancel

## 9. Database lookup (Database connection, Database explorer)







Database lookup

Step name: Database lookup

Connection: MySQL-ClassicModels3 Edit... New... Wizard...

Lookup schema: classicmodels Browse...

Lookup table: customers Browse...

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	customerN...	=	custo...	

Values to return from the lookup table :

#	Field	New name	Default	Type
1	custo...			String

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

Help OK Cancel Get Fields Get lookup fields

8. Select values (Get fields to select, Get fields to change).





Select values

Step name Select values

Select & Alter Remove Meta-data

Fields :

#	Fieldname	Rename to	Length	Precision
1	orderNumber			
2	customerNumber			
3	Total			
4	customerName			

Get fields to select

Edit Mapping

Include unspecified fields, ordered by name ☐

Help OK Cancel

Select values

Step name Select values

Select & Alter Remove Meta-data

Fields to alter the meta-data for :

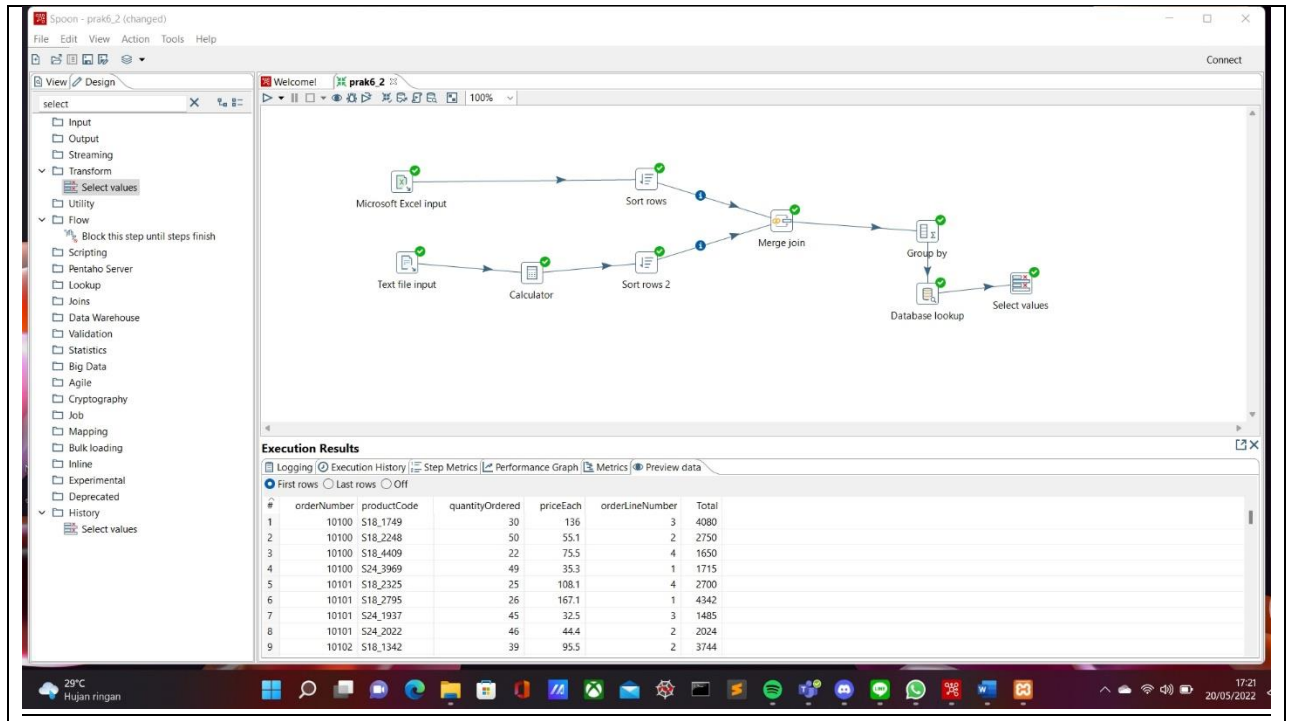
#	Fieldname	Rename to	Type	Length	Precision	Binary to Normal?
1	Total					N
2						
3						
4						

Get fields to change

Help OK Cancel

## 9. Output





#### 4. File Praktikum

Github Repository:

<https://github.com/IbnuFajar7/Data-Warehouse/tree/main/Prak-6>

#### 5. Soal Latihan

Soal:

1. Apa yang dimaksud dengan Merge Data?
2. Apa tujuan dan fungsi dari menggabungkan data?

Jawaban:

1. Menggabungkan Data
2. Tujuan dan fungsi dari menggabungkan data adalah untuk memudahkan kita ketika harus memindahkan data 1 dan yang lainnya, fungsi nya adalah lebih memudahkan melihat perbandingan data 1 dan yang lainnya ketika ada table yang sama, hitungan yang sama.

#### 6. Kesimpulan

- a. Dalam pengerjaan praktikum Data Warehouse, kita harus benar-benar teliti dalam menginputkan suatu fungsi untuk menampilkan suatu keluaran pada layar dengan sesuai.
- b. Kita dapat mengetahui bahwa pentingnya untuk melakukan penggabungan data adalah sangat penting, metode ini sangat berguna untuk kita agar bisa melihat dan membaca suatu data yang berbeda lebih mudah untuk kedepannya.

#### 7. Cek List (✓)



Jurusan Teknik Informatika & Sistem Informasi  
Fakultas Teknologi Industri – Universitas Trisakti

8.

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	✓	
2.	Latihan Kedua	✓	

## 9. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	20 Menit	1
2.	Latihan Kedua	20 Menit	1

Keterangan:

1. Menarik
2. Baik
3. Cukup
4. Kurang

